

Weekly Report

Yuxin Ma

2017.06.26 - 2017.07.02

Projects

InfoVis Paper Review

According to the review plan, I have re-organized the machine learning sections. Other points in the plan will be done in the next week.

Misc.

- Framework of the slides is done, which only needs filling some examples.

Plan for the Next Week

- Slides for the summer school on Monday.
- InfoVis paper review: including details in the labeling sections and the user study sections.

Paper Reading

Two papers [1, 2] related to predictive visual analytics were proposed in EuroVis 2017 STAR.

The State of the Art in Integrating Machine Learning into Visual Analytics

Enrico Bertini are the original authors of their first survey paper in 2009 [3]. However this time it changes to Alex Endert. This survey gives a comprehensive summarization and taxonomy of how visualization helps machine learning approaches, and vice versa. In the discussion of remained challenges, they do mentioned how to understand users' interaction with machine learning approaches, which is related to prescriptive visual analysis.

The State-of-the-Art in Predictive Visual Analytics

Coming from Michael Gleicher and Ross, this paper gives a formal definition and pipeline of predictive visual analytics (PVA). In general, predictive visual analytics uses predictive models (commonly referred to machine learning models) as the automated analysis method in the

visual analytics loop. The users are intended to process related training data, analyze features, adjust models and validate results of the predictive model as well as gain information and extract knowledges from the entire process. Figure 1 compares the differences of goals between automated predictive analysis method and predictive visual analytics methods.

Table 2: *The goal of each step in the PVA pipeline and the general predictive analytics procedure.*

	PA	PVA Exclusive
Overall Goal	Make Prediction	Support Explanation
Data Preprocessing	Clean and format data	Summarize and overview the training data
Feature Selection and Generation	Optimize prediction accuracy	Support reasoning and domain knowledge integration
Modeling	Optimize prediction accuracy	Support reasoning and domain knowledge integration
Result Exploration and Model Selection	Model quality analysis	Get insights; Select the proper model; Feedback for model updates
Validation	Test for overfitting	Get insights from other datasets

Figure 1: Comparison of goals between predictive analysis and predictive visual analytics.

Several challenges should be addressed:

- Knowledge generation and integration in PVA. Currently there is no such a framework which explains how the knowledge can be generated in the predictive visual analysis process.
- Interpretable/Explainable AI, which is currently a research direction in the field of AI. It requires the AI methods to take interpret-ability and explain-ability into consideration in the algorithm design.

References

- [1] Y. Lu, R. Garcia, B. Hansen, M. Gleicher, and R. Maciejewski, “The State-of-the-Art in Predictive Visual Analytics,” *Computer Graphics Forum*, vol. 36, no. 3, 2017.
- [2] A. Endert, W. Ribarsky, C. Turkay, B. L. W. Wong, I. Nabney, I. D. Blanco, and F. Rossi, “The State of the Art in Integrating Machine Learning into Visual Analytics,” *Computer Graphics Forum*, vol. 00, no. 00, pp. 1–28, 2017.
- [3] E. Bertini and D. Lalanne, “Surveying the complementary role of automatic data analysis and visualization in knowledge discovery,” *Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery Integrating Automated Analysis with Interactive Exploration - KDD '09*, pp. 12–20, 2009.